REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-01-0188

1. REPORT DATE *(DD-MM-YYYY)*
18-10-2005

2. REPORT TYPE
Conference Proceedings

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
Modeling and Simulation for the Knowledge Management for Distributed Tracking (KMDT) Program

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHORS
Marion G. Ceruti
Tedd L. Wright
Dwight R. Wilcox
Scott C. McGirr

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
SSC San Diego
53560 Hull St.
San Diego, CA 92152-5001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
This is the work of the United States Government and therefore is not copyrighted. This work may be copied and disseminated without restriction. Many SSC San Diego public release documents are available in electronic format at: http://www.spawar.navy.mil/sti/publications/pubs/index.html

14. ABSTRACT
This paper describes the application and the approach to modeling and simulation for Knowledge Management for Distributed Tracking (KMDT). This is an ongoing research and development program to explore methods to improve command, control, and decision support functions in the battle space. The focus of the simulation effort is on a hypothetical scenario designed to simulate how knowledge management technologies, such as ontologies and intelligent agents can be used to improve battle space awareness. New decision-making processes are needed in command centers to enable distributed network-based tracking. Agents can use web services to access data and schemas at multiple platforms in the battle space. New methods are needed to support distributed tracking with dissimilar sensors such as RADAR and SONAR. These approaches can reduce the uncertainty in the detection, localization, classification and identification of unknown contacts.

Published in *Proceedings of the International Workshop on Modeling & Applied Simulation, MAS 2005*, 16-18 October 2005, Bergeggi, Italy.

15. SUBJECT TERMS

Autonomous agents	modeling and simulation	sensors
decision support	ontology integration	tracking
knowledge management	ontology methodology	

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19B. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	9	Marion Ceruti, Code 246206 / (619) 553-4068

The International Workshop on

Modelling & Applied Simulation

MAS 2005

EDITED BY

Agostino Bruzzone
Roberto Mosca
Chiara Briano
Edward Williams

October 16-18 2005
Bergeggi, Italian Riviera, Italy
Congress Center & Hotel "Claudio"

Printed in Genova, Italy, September 2005
ISBN 88-900732-8-4

Table of Contents

Modeling and Simulation for the Knowledge Management for Distributed Tracking (KMDT) Program

Marion G. Ceruti,* Tedd L. Wright, Dwight R. Wilcox, and Scott C. McGirr
Space and Naval Warfare Systems Center, San Diego (SSC-SD), Code 246206
53560 Hull Street, San Diego, CA 92152-5001, Tel. 619 553-4068*
marion.ceruti@navy.mil, tedd.wright@navy.mil, dwight.wilcox@navy.mil, scott.mcgirr@navy.mil

Abstract

This paper describes the application and the approach to modeling and simulation for Knowledge Management for Distributed Tracking (KMDT). This is an ongoing research and development program to explore methods to improve command, control, and decision support functions in the battle space. The focus of the simulation effort is on a hypothetical scenario designed to simulate how knowledge management technologies, such as ontologies and intelligent agents, can be used to improve battle space awareness. New decision-making processes are needed in command centers to enable distributed network-based tracking. Agents can use web services to access data and schemas at multiple platforms in the battle space. New methods are needed to support distributed tracking with dissimilar sensors such as RADAR and SONAR. These approaches can reduce the uncertainty in the detection, localization, classification and identification of unknown contacts.

Keywords: Autonomous agents, decision support, knowledge management, modeling and simulation, ontology integration, ontology methodology, sensors, tracking

1. Introduction

New approaches to military Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance (C4ISR) using knowledge management technologies such as ontologies [6] and intelligent agents [5] [8] are being explored. These new approaches are intended to help make FORCEnet [8] [10], the U.S. Navy's operational construct and architectural framework for naval warfare in the information age, a reality. In particular, these new technologies may revolutionize traditional detection, localization, and tracking systems used in undersea, surface, and air warfare. Unlike older legacy systems that primarily rely on similar sensor types to detect, localize, and track Targets Of Interest (TOIs), these new approaches can also use dissimilar sensor types to participate in level-one data fusion tasks (that is, detection, localization, classification, and identification) [19], [25]. During task execution, intelligent agents can access sensor ontologies to identify relevant sensor data meeting current requirements. Guided by these ontologies, the agents can then access sensor platforms in the battle space via web portals hosted on secure communications networks. Each platform or shore-based sensor station will have a web portal. Each web portal will have multiple pages, one for each sensor at that site. Te sensors can report multiple contacts. Fusion of these data, either at distributed sites or at a command center, will help reduce uncertainty in the battle space.

The goal of KMDT is to allow war fighters to reduce uncertainty by better organizing and using the data collected from existing sensors. In order to achieve this goal, KMDT will initially focus on technologies that are essential for the design of next-generation tracking systems that use knowledge management techniques, and network-based command and control. A Modeling and Simulation (M&S) approach will be taken using a hypothetical scenario in order to develop and evaluate new knowledge management techniques [22]. M&S of information flow in the battle space is a relatively inexpensive way to depict both baseline use and more efficient future uses of existing sensors and their data output, without costly field trials. The ability to run multiple trials using an M&S approach facilitates the generation of statistics useful in evaluating the effect of fused information on the reduction of uncertainty in the battle space.

This paper is organized as follows. Section 2 provides background information on KMDT and discusses the motivation for the program. Section 3 describes a concept of operations designed to show how the technology could be employed in a battle-space environment. Section 4 presents an outline of the modeling and simulation method for KMDT. Section 5 discusses ontologies and their integration for KMDT. Section 6 describes metrics and simple statistics for evaluating and documenting the simulations. Lastly, section 7 describes directions for future research.

2. Background and Motivation for KMDT

Sensors deployed from mobile platforms such as ships and aircraft, and fixed platforms such as ground-based stations, can provide both passive and active information on unknown contacts and potential TOIs in their vicinity. Passive information is derived from signals generated by the contact/target that propagate to the sensor, while active information is derived from signals originating at the sensor system that propagate to the contact/target and generate a return. Typically, passive signals provide a greater detection range since the signal, which attenuates as it propagates, travels only in one direction. Spatial processing of passive signals from sensors can provide Line of Bearing (LOB) information, and temporal processing can often provide frequency attributes of the signal that can help classify/identify the contact. Active signals, on the other hand, while suffering two-way propagation loss, can provide an estimation of the range to the contact by measuring the travel time of the signal. Furthermore, characteristics of the signal, such as signal strength and waveform, can be controlled to provide designed responses from the contact. In addition to LOB and range information, passive and active sensors can provide such information as velocity and acceleration, size or orientation, as well as classification and identification signatures of the contact or target.

KMDT will initially focus on fusion of LOBs to potential contacts via cross fixing, which is using the intersection of LOBs from two or more platforms to localize contacts. This LOB cross fixing can in principle involve either similar (homogeneous) and/or dissimilar (heterogeneous) sensors. Homogenous sensors are of the same type, measuring essentially the same physical parameters of the unknown contact. For example, if both sensors are passive acoustic sensors, the same signal type, in this case acoustic, is obtained. Similarities or differences are readily observed based on a comparison of the signals, thus facilitating subsequent data fusion and potential association of the signals as emanating from the same contact.

In contrast, heterogeneous sensors are of different types, such as acoustic and electromagnetic sensors. Signals derived from these sensors seldom look the same even if they represent the same contact. Similarities and differences must be determined indirectly, unlike the case of homogeneous sensors. Therefore, it is reasonable to conclude that LOB cross fixing is inherently more difficult with heterogeneous sensors than with homogeneous sensors. In the heterogeneous case, each signal must be analyzed separately to determine the set of potential contacts that could have produced the signal. An unknown target that could have produced both signals will then likely be a member of the intersection of the sets of potential contacts.

Cross LOBs from homogeneous sensor data are used routinely in ship and aircraft navigation to determine position. However, the use of heterogeneous sensor data to determine the position, classification and identity of unknown contacts and potential targets in the battle space has not been utilized effectively. LOB information from heterogeneous systems that could potentially help to localize a contact often does not reach a command center and contribute to the decision process at all. Sometimes the heterogeneous data cannot be transmitted efficiently or there is no perceived payoff for their propagation. Even when the data are available, they may not be fused with existing data because they are unfamiliar to operators or because they appear too dissimilar, incomplete, or fragmented to correlate with existing data. In addition, due to separate processing time scales, and transmission requirements, the data may not be available in a timely manner.

Currently, command center personnel often are overloaded with tasks and uncorrelated information. Conversely, they sometimes have difficulty in obtaining information that they need to confirm decisions in a timely manner. Often decisions are made using uncertain information. Uncertainty, in turn, contributes to battle stress. Intelligent agents can relieve overloaded operators by retrieving more complete information from existing sources. The availability of this additional information in the battle space is aimed at reducing tracking uncertainty and targeting errors.

3. Concept of Operations

This section describes a concept of operations (CONOPS) that shows capabilities under development in KMDT that will be used in future command and control operations. It forms the basis of a scenario for the modeling and simulation effort. It is assumed that CONOPS will evolve from existing Tactics Techniques and Procedures (TTPs).

Figure 1 illustrates an example CONOPS. A commander on board northbound ship A receives a RADAR report of an unknown contact detected at a bearing of 045 degrees. Although the contact cannot be classified or localized using only the information in the report, it is believed that the contact is a surface vessel. Since the commander does not know whether the contact corresponds to a potential threat, he orders an operator to obtain more information. The operator tasks an intelligent agent to search a Local Area Network (LAN) for friendly platforms in appropriate sectors of the battle space that correspond to zones situated NW and SE of the unknown contact. The definition of these sectors can either be under the control of the operator or the intelligent agent, and can be changed depending on search results. In this example, no friendly platform is identified in the SE sector, but friendly Ship B is available in the NW sector.

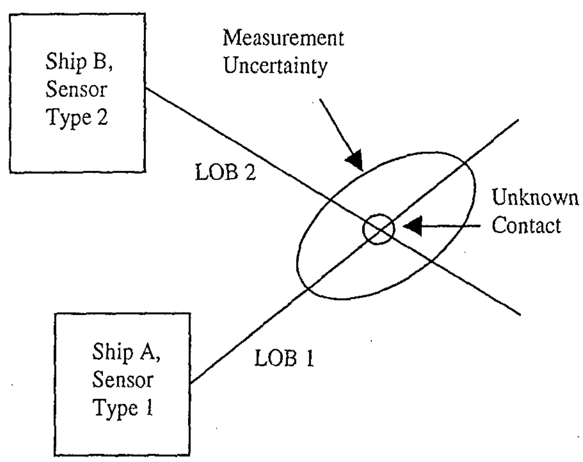

Figure 1. Detection geometry showing lines of bearing from ships A and B detecting an unknown contact with heterogeneous sensor types 1 and 2.

After determining the availability of friendly platforms in appropriate sectors, the agent decides whether or not the data on those platforms can be useful to the commander on Ship A. Using the secure LAN, the agent finds the sensor-ontology web portal to correlate the known capabilities of the platforms with the kinds of information that could be combined with the RADAR of Ship A to help classify and localize the unknown contact. In this example, Ship B is determined to have acoustic sensors that can potentially provide LOBs and other acoustic signature information.

An agent is then tasked to collect data from the web portal of Ship B. Using knowledge acquired from the sensor ontology, the agent extracts appropriate data from the web portal that meet certain correlation and satisfaction criteria. For example in Figure 1, the agent determines whether or not Ship B has posted acoustic LOB(s) that could intersect the contact LOB from Ship A. If so, the agent collects the appropriate data, along with the date-time group of the measurements with respect to Ship B. The agent from Ship A can also direct another agent on Ship B to search for other available information regarding the unknown contact, such as any reports from friendly aircraft in the area, or to query its sensors for new contacts. This information is all posted to the web portal of Ship A. An intelligent agent on Ship A issues an alert to the operator's workstation on Ship A, that a report from Ship B is available on the LAN. The operator fuses this information with the RADAR contact from ship A and recommends a classification (hostile, friendly or neutral) of the unknown contact to the commander who now has enough information to take action.

Sometimes an initial area search for participating platforms identifies more than one candidate platform that observed a useful LOB. In this case, the agent can acquire additional potential LOBs and send messages to Ship A containing sensor data from the ships collecting these data or from the shore-based sensor stations. Alternately, the commander on Ship A can specify additional constraints to restrict the search space of the intelligent agent to minimize data overload on the part of the analysts. Finally, if the agent finds no platforms that have potentially useful LOBs a message can be sent indicating that the search has concluded with a negative result.

As illustrated in the CONOPS above, the complexities in KMDT include dealing with potentially large amounts of information (both positive and negative), and alignment of sensors in a common time and space frame of reference. Data from sensors can include LOB, range, velocity from Doppler radar, acceleration, pulse repetition rate, peak frequency, etc., some of which are listed in Table 1. In addition, a priority for use of the sensor data in the fusion process may be established; for example, first correlating data from homogeneous sensor types, followed by correlating data from heterogeneous sensors, and finally, correlating data from dissimilar sources (e.g. ships and ground stations) and according to combat identification.

In Table 1, the data source could be a web portal, a message, a visual observation, etc. Mostly in this simulation the source will be a web portal of a ship or a shore-based sensor station on the secure network. The data for "reference files" in Table 1 could be a photo, a lofargram (in the case of passive acoustic sensors), a radar image, or some other supporting file with multimedia data that may be of use in the level-one fusion task.

Table 1. Sample intelligent-agent sensor-association table for simulation where AOU means "area of uncertainty."

Output page Web data	Sensor 0	Sensor 1	Sensor 2	Sensor 3
Platform type	CG88			
Platform Name	Own Ship	Ship B	Ship C	MMA
Platform ID	SN332			
Operator				
Platform Lat				
Platform Lon				
Geographic Correlation		85	90	---
Observation Date-Time Group	1 Apr 05 10:21:45			
LOB number		LOB 2	LOB 3	
Contact LOB		130°	332°	
LOB AOU major axis		8	4	
LOB AOU minor axis		5	3	
AOU Major-axis orientation		Parallel to LOB	Perpen-dicular to LOB	
Sensor type		EM	SONAR	ELINT
Sensor mode		Passive	Active	Passive
Sensor range		15 NM	10 km	
Sensor Coverage			Side oriented	Tasking required
Sensor Lat				
Sensor Lon				
Sensor error		1 NM	1 Km	
Attribute Correlation		90	85	---
Reference Files				

4. Modeling and Simulation

The KMDT approach to target detection, localization, classification, and tracking will use the results of agent-based data fusion to reduce uncertainty and improve command decision efficiency. This approach relies upon the availability of appropriate information content and flow through the battle space. This information will be represented in messages posted to web portals on individual platforms and accessible to agents over a secure Local Area Network (LAN). The agents will access the sensor ontology to define data requirements and sensor capabilities, and improve message content comprehensiveness.

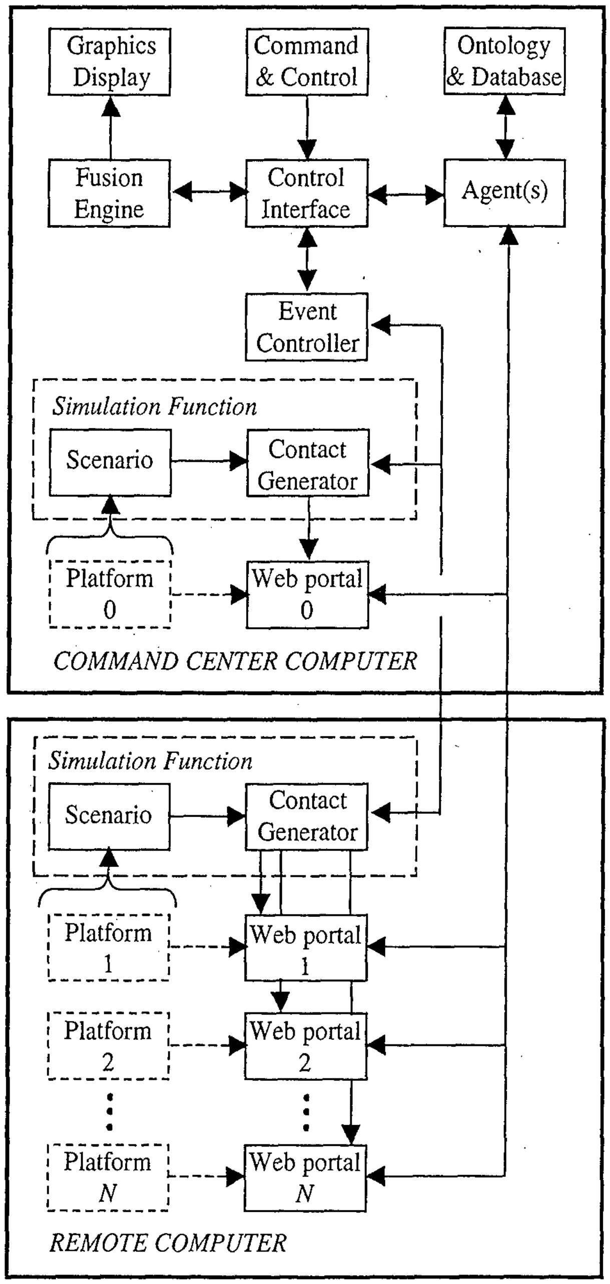

Figure 2. KMDT Simulation Design Diagram

To demonstrate feasibility of concept, a Modeling and Simulation (M&S) effort will be conducted. The initial M&S effort will focus on a simplified concept of operations and scenario. The messages posted to the web portals will consist primarily of LOB information derived from sensors of various types (active or passive electromagnetic, acoustic, or optical), upon which some analysis has been performed in order to reduce "raw" sensor outputs to metadata contact information. Not only would unprocessed sensor data be difficult for agents to interpret, but such data would also increase bandwidth requirements over the LAN. For the simulation, each sensor platform has a standard web portal such as shown in Table 1, the URL of which is known to all other platforms or stations that are on the LAN. When more information is desired about an unknown contact, an intelligent agent is deployed that finds pertinent web portals, reads data on these web portals, and evaluates whether the data on the pages satisfy a pre-determined set of criteria. Initially, the approach will be for agents to perform passive query of web portals that does not involve platform response. Platforms simply update their web portals when new data are received and processed at the message level. Sensor data posted on the web portal will have an associated date and time. Each mobile platform updates its position on the web portal frequently. The frequency of update can be controlled in the simulation.

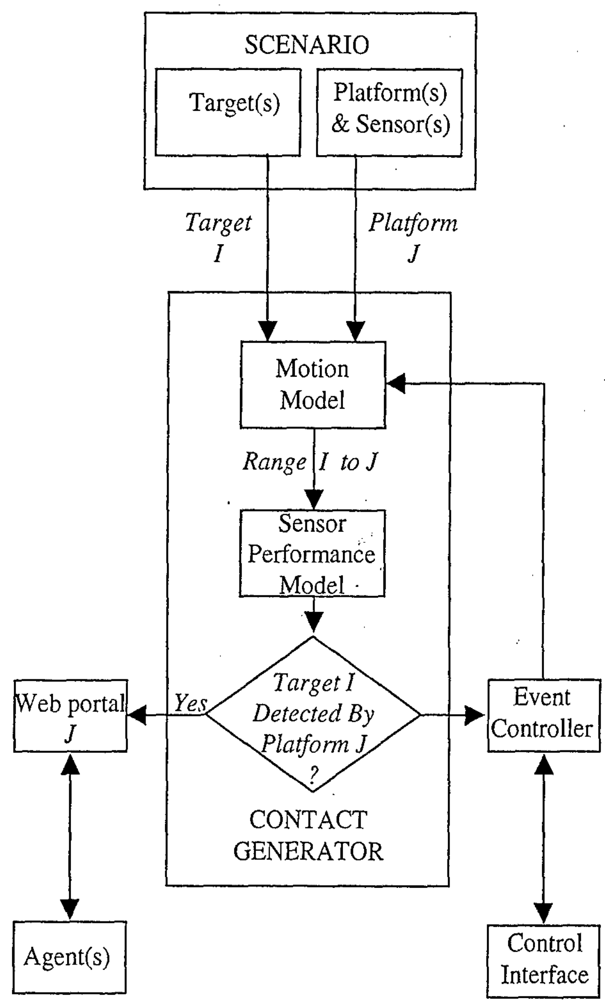

Figure 3. KMDT Simulation Functional Diagram.

The initial simulation design is illustrated in Figure 2. The simulation consists of two computers connected over a LAN. The "Command Center Computer" provides control and input/output functions for the simulation, including the tasking of agents to gather both organic (own platform) and inorganic (remote platform) data for fusion via the fusion engine. In addition, this computer runs simulation models, simulates organic sensor data through an event-controlled scenario script, and posts organic contact data on a web portal.

The "Remote Computer" simulates inorganic data from multiple platforms through an event controlled scenario script and posts the data on separate web portals representing each separate platform. That way, a single computer can be used to simulate multiple platforms that would ordinarily be distributed in the battle space. Despite the use of a single computer, the scenario will allow for independent operation of the remote platforms under the control of an Event Controller, that synchronizes the actions of the platforms and target(s) in the simulation.

The roles of the Scenario and Contact Generator modules in the simulation are detailed in Figure 3. The Scenario will consist of predefined platforms in some theater of operations, with specified capabilities, sensors, and initial locations. The capabilities and performance parameters of the potential sensors will be defined in a separate database and associated ontology. Additionally, the Scenario will contain one or more targets with specified capabilities and initial locations.

As the simulation progresses under the control of an Event Controller, the positions of the target(s) and mobile sensor platforms will be updated at each step of the simulation according to motion models appropriate for the target(s) and platforms contained in the Motion Model module. These motion models may consist of predefined tracks or randomly determined motions subject to constraints. In order to support the detection decision within the Contact Generator, the range and LOB between each target I and platform J will be computed. The Sensor Performance Model then determines the detection range of each sensor on each platform, and this range is compared against the distance between the sensor platform and each target. If it is determined that a target is within the detection range of a sensor on some platform, LOB information is posted on that platform's web portal.

The simulation is modularly designed to enable different sub-models with varying fidelities to be evaluated. For example, the first simulation might employ static platforms and constant environments, wherein the sensor performance sub-models are characterized by simple "cookie cutter" constant detection ranges. As proof of concept is established, more realistic range and environment dependent sub-models can be substituted.

Eventually, as proof of concept is simulated, *distributed fusion* concepts will be examined. With distributed fusion, the interaction of agents with web portals will be accomplished via a series of active "query" and "reply" messages. In response to a query the initial fusion takes place on the platform before replying to the location where the query originated. Subsequent fusion of information from multiple platforms occurs at the location of the initial query.

The functional components of the KMDT approach and their interrelationships are shown in Figure 4. Each participating sensor system that supports a web page interface reports tracks detected by the sensor. The content and format provided by these web pages vary depending on the nature and capabilities of the respective sensor.

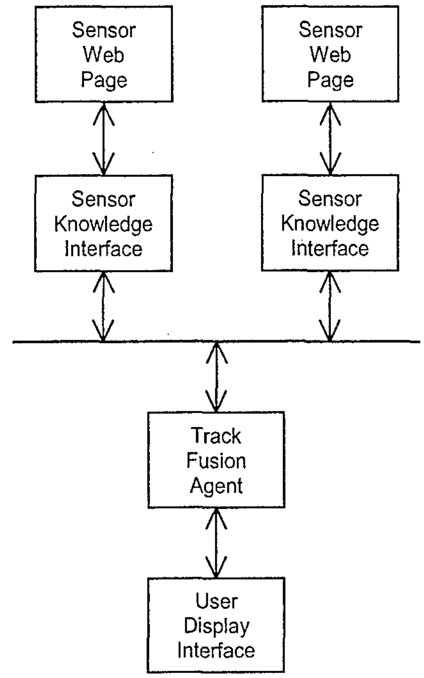

Fig. 4. KMDT Application Functional Diagram.

The sensor web pates are asynchronously read by the Sensor Knowledge Interface(SKI) Associated with the respective sensor. The SKI's response to queries they receive from external intelligent agents based upon knowledge they capture by reading their associated sensor web pages. The track-fusion agent generates query to and accepts the associated responses from the SKIs. It processes the query responses to suggest likely associations among the detection reports on the various sensors and displays the results for the human operator.

The sensor web pages generally provide information relative to the sensor platform position. To provide the sensor ontology a common frame of geographic reference frame for the region of interest, all agent queries and SKI responses include definition of a polygon with vertices defined by latitude and longitude. The size and shape of the polygons n the response depend on the resolution of the particular sensor. Sensors that report both range and bearing provide a small quadrilateral indicating where the track was detected. In contrast, sensors that detect only bearing provide a narrow pie-shaped region originating at the sensor's position and truncated by the range detection limit. The fusion process consists of calculating where the polygons from the various sensors intersect in combination with other non-spatial yet potentially compatible factors.

5. Ontology Identification and Integration

Knowledge and concept representations accessible to agents on the network form the basis of the ontology for information sharing and automated processing. Access to the common, integrated ontology will enable agents to coordinate their interactions with each other and with operators. Agents can coordinate message-level data retrieval, fusion, and integration, thus bringing key missing data to the attention of decision makers in command centers. This intelligent-agent based method will exploit distributed data now unused for LOB tracking.

Different sensor communities have evolved their own concepts, terminology and procedures. For some of these communities, ontologies have been developed that describe some sensor-related concepts. However, a comprehensive sensor ontology is needed to provide a common understanding about sensors and the aggregate of their data. A common sensor ontology defines concepts and map metadata from disparate systems and communities. This ontology supports intelligent agents and also can serve as the basis for knowledge-base development and intelligent data fusion. Knowledge and concept representations accessible on the network form the basis of the ontology for information sharing and automated processing.

Various approaches will yield a common ontology. One way is to use a common ontology and data-reference standard such as the Command and Control Core Data Model [16], [17] or the Command and Control Information Exchange Data Model [4]. The Center for National Research Initiatives has advocated a common-metadata approach [3], which is used by the Library of Congress and US patents office. Another approach is to map ontological information (e.g. concepts and relationships) from schemas in distributed sources utilizing namespaces and protocols to facilitate the procedure. The best approach may be a combination of these two approaches.

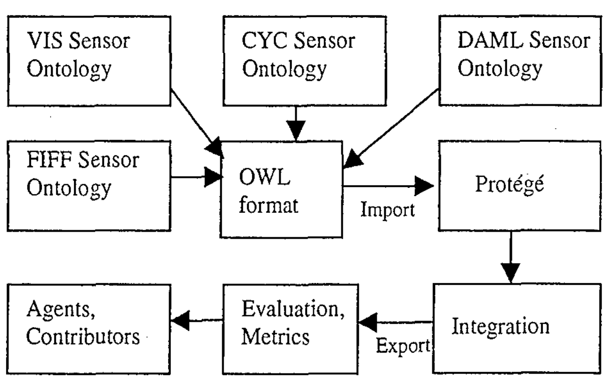

Fig. 5. Ontology collection, standardization, integration, evaluation and utilization in the KMDT program.

Ontology development for KMDT is in progress using Protégé [14], an ontology and knowledge-base development tool, and OWL [14], [20], a web-ontology language. Protégé [14] was selected as the knowledge-based system development tool because of its capabilities, user-friendly interface, documentation, tutorial availability, and large user base. Protégé and OWL support the Resource-Description Framework (RDF) format for schema and file sharing [14].

A survey of existing sensor ontologies was performed. Figure 5 summarizes some mappings for the ontologies found in the survey. These sensor ontologies together with selected concepts are listed in Table 2. Several ontologies were found but none was complete and no two ontologies were exactly alike. No single ontology included all of the concepts in sensor data acquisition, fusion, interpretation, and usage, nor did any of the existing ontologies include all of the concepts of any of the other sensor ontologies.

Table 2. Some noun concepts represented in various sensor ontologies where "x" means explicit representation and "i" means that the concept is implicit as instance or related concepts.

Ontology / Concept	VIS [15]	Cyc [24]	DAML [21]	FIFF [12]	S.Davis [11]
Sensor type:	x	x	x	x	x
Active	x	x	x	x	x
Passive		x	x	x	x
Radar		x		x	x
Acoustic	x	x			
Magnetic	x	x			
Electro-optic	x	i		x	
ElectroMagnetic		x	x	i	x
Mechanical		i	x		
Biological		x	x		
Chemical		x	x		
Radioactive			x		
Cyber		i	x		
Optical	i	x	x		
Microwave				x	
Geometry	x	x		x	
Track	x	i			
Signal	x	x	i		x
Resolution			x		
Environment	x	x			
Target (contact)	x		x	x	
Range		x	x	x	
Goal		x	i	x	
Sensor mode			x		
Sensor location	x	x	x		i
Sensor data	x	x	x		

For example, Versatile Information Systems, Inc. (VIS) under contract to the Office of Naval Research and in collaboration with the Space and Naval Warfare Systems Center has developed several related ontologies, including a pedigree ontology describing the concepts about single sensors [15]. Cycorp, which has developed a large knowledge base called "Cyc," has dedicated a section of the Cyc ontology to sensor concepts. (See, for example, [24]). J. Hendler of the University of Maryland and co-workers have developed a sensor ontology in the Defense Advanced Research Projects Agency, (DARPA) Agent Markup Language (DAML) [13], [20], [20].

The Formal Information Fusion Framework (FIFF) ontology models the multi-sensor application domain, including data-fusion theory [12]. However, [11] is focused in different areas of the sensor-ontology domain, citing for example, specific instances of satellite sensors. Almost all sensor ontologies include the concepts of "active" and "passive" sensors but at different levels. For example, in Cyc, the concepts occur as specializations of SONAR, whereas in [12] they are specialized at higher level, under "sensor." In [11] "active" is specialized between "group-based sensor" and "radar," and "passive" is specialized between "space-based sensor" and "infrared sensor."

In general, ontology integration involves mapping of concepts between different ontologies. The terminology, structure and representations of ontologies can differ in many ways [24]. For example, concepts can have different names in different ontologies or they can occur at different levels of specialization/generalization [24]. These differences also are found in databases integration [7]. Also, ontologies can be disjoint or one can be a subset of the other [24]. Ontologies can be organized in totally different fundamental representations. For example, a probabilistic ontology may form the basis of a knowledge base that is organized as a Bayesian network whereas a deterministic ontology may form the basis of a knowledge base consisting of rules and axioms in a truth-logic system.

6. Metrics and Statistics

6.1 Agent Metrics

Metrics and statistics can be used to evaluate and document the behavior of agents in simulations. The following metrics can track the activity of the agents [9]:

1. The number of web portals accessed to search for the desired data,
2. The number of relevant data retrieved from each site,
3. The number of successful data retrievals vs. the number of agent deployments on an individual-agent basis,
4. Same statistics as in item 3 above, but summarized to include all agents deployed during a given time period in the simulation.

To analyze agent errors, the following metrics can be collected [9]:

1. The number of irrelevant data retrieved from each site,
2. The number of incorrect data retrieved from each site,
3. The number of correct data that the agents could have retrieved but did not (a manual analysis that involves keeping track of all possible alternatives in the simulation and comparing the results to the alternatives.)

One way to collect the data on agent errors is to save the history of the simulation scenario, including the distribution of platforms, in a log file for later analysis [9].

6.2 Ontology Metrics

Ontology metrics can be used in a variety of integration applications. For example, they can be applied to a common ontology reference prior to processing and integration, or they can be applied to schema matching in eXtensible Markup Language (XML) integration. (See, for example, [23] and [18].)

General statistics - The development of an integrated sensor ontology can be tracked with simple metrics. One metric is the number of initial concepts input into Protégé/OWL. Some other metrics associated with concept acquisition are 1) the number of added ontologies; 2) the total number of concepts in the proposed ontology prior to integration; and 3) the number remaining in the integrated ontology, assuming all non-redundant concepts are retained. Metrics associated with ontology integration are 1) the number of redundant concepts deleted because they were not needed; 2) the number of concepts added to fill gaps that became apparent during the integration process; and 3) the number of remaining concepts in the final integrated ontology. Still another dimension of metrics is to count the number of levels in the ontology hierarchy and the classes or instances residing at each level.

Individual disjunction metric - In addition to the metrics described above, a method is needed to characterize, estimate, and eventually measure disjunction in information systems, and particularly in ontology-integration tasks. Class cohesion has been studied in object-oriented systems and metrics have been developed [1], [2]. Ontologies are hierarchical structures similar to structures in object-oriented systems. The cohesion metrics measure cohesion between members of the same class whereas the disjunction metric described below tracks the placement of the same concept in related ontologies.

A disjunction metric proposed here specifies the degree of disjunction in ontologies by identifying the level of generality or specificity at which a concept occurs in one ontology, compared to the level of occurrence in another ontology. The disjunction metric is useful in an ontology-integration application when comparing the value added of various ontologies that were developed separately from different sources.

To apply the metric (Dj in equation (1) below), all levels in the hierarchy of concepts in each ontology must be labeled with 1 representing the most specific instances, and higher numbers representing upper-level ontologies.

(1) $Dj\ (O_1(c_i), O_2(c_k) \ldots O_p(c_m)) = (i, k, \ldots m)$

Equation (1) defines the disjunction metric, Dj as a set of levels at which a common concept occurs in a collection of ontologies. In (1), "c" is a concept that occurs at level "i" in ontology 1, which is called "O_1." The same concept, c, occurs in ontology 2, called "O_2," at level "k." Concept "c" also occurs at some arbitrary level "m" in ontology p, called "O_p." The "..." in (1) means that the number of ontologies that can be compared in this manner is not restriction. For example, equation (2) below illustrates the disjunction metric in an hypothetical case of two ontologies, 1 and 2. If common concept "c" found at level 3 in ontology 1, were also found at level 5, in ontology 2, one could write the disjunction metric as follows:

(2) $Dj\ (O_1(c_3), O_2(c_5)) = (3, 5)$

Overall disjunction metric – Equation (1) is meant to express disjunction for a single concept. However, many concepts are found in any meaningful ontology. To measure and compare the characteristics of various ontologies, an overall disjunction metric is needed that includes multiple concepts, not just one. To calculate an overall estimate of disjunction, each index (i, k, ... m) can be averaged separately across a group of concepts that occur in the same ontology. An overall disjunction metric for two ontologies can be calculated based on average values of the levels for a collection of "n" concepts:

(3) $< Dj\ (O_1, O_2) > = (\Sigma\, i/n,\ \Sigma\, k/n)$

where the instances of i and k are the values of each pair of levels found for each common concept. To use this metric, the ontology that pertains to each knowledge base (KB) must be sufficiently complete to locate the corresponding levels in the ontologies. Another way to conceptualize the disjunction metric in (2) is to consider that a concept at level 3 of ontology, O_1, is equivalent to a corresponding concept at level 5 of ontology, O_2. The usefulness of disjunction metrics will increase when a more standardized way to organize an ontology is developed. Dj and <Dj> will depend not only on concepts in common but also on the structure of the various ontologies.

For example, consider an overall disjunction metric based on equation (2) and two others like it so that "n" in (3) is 3:

(2) $Dj\ (O_1(c_3), O_2(c_5)) = (3, 5)$
(4) $Dj\ (O_1(c_2), O_2(c_4)) = (2, 4)$
(5) $Dj\ (O_1(c_1), O_2(c_3)) = (1, 3)$
(6) $< Dj\ (O_1, O_2) > = ((3+2+1)/3,\ (5+4+3)/3)$
(7) $< Dj\ (O_1, O_2) > = (2, 4)$

An assumption in equations (2) through (4) is that the each equation addresses a distinct concept. If the overall disjunction metrics are low (e.g. (1,2)) it indicates that the ontologies have common concepts at the same level of granularity, which is important to know for integration purposes. It is also a signal to look for duplicate concepts and delete any redundant information. If the metrics are high, (e.g. 5,7) it implies that: 1) The ontologies proposed for integration may contain concepts that have been overlooked and therefore are missing in the integrated ontology, or 2) The proposed addition to the ontology has nothing to do with the main topic.

Application example – To apply the individual disjunction metric, Dj, consider for example the concept, sensor, which occurs in all ontologies listed in Table 2, as it is the most general. Dj is given by equation (8).

(8) $Dj\ (O_{VIS}(sensor), O_{CYC}(sensor), O_{DAML}(sensor), O_{FIFF}(sensor), O_{DAVIS}(sensor)) = (5, 5, 6, 4, 7)$

Dj will depend on the structure of the various ontologies and therefore can provide at a glance some insight regarding the relative ontology hierarchies. Low numbers for very general concepts, such as 1 or 2, indicate a very flat ontology whereas higher numbers, such as the ones in (4), indicate more levels of specialization/generalization.

7. Directions for Future Research

A further refinement to this study would be to account for the error in measured data, such as positions, frequencies, etc. by representing these data not as fixed points but as probability distributions so that principles of fuzzy logic can be applied, thus providing a more realistic simulation.

Using the statistics gathered from this study, described in section 6, the simulation can be refined further to address any noted anomalies in the statistical data.

The present design, for which agents acquire message-level data from remote web portals to be fused at the site that deployed the agents, utilizes a centralized fusion architecture. In contrast, a future design could require the agents to process the message-level data from the web portals remotely at the site from which the data are retrieved. This distributed architecture design possibly can relieve overloaded operators tracking multiple unknown contacts and who may have deployed several agents to retrieve data on each one.

More advanced agent capabilities can be developed to include semantic understanding of the content of web portals. This is possible with description and discovery protocols on open service-oriented architecture. One output from the agent can be a track declaration from associating and correlating multiple LOBs.

Acknowledgements

The authors thank the Office of Naval Research and the Space and Naval Warfare Systems Center, San Diego, Science and Technology Initiative for their support of this work. This paper is the work of U.S. Government employees performed in the course of employment and no copyright subsists therein. It is approved for public release with an unlimited distribution.

References

[1] L. Badri and M. Badri, "A Proposal of a New Class Cohesion Criterion: An Empirical Study," *Journal of Object Technology*, pp. 145-159, vol. 3, no. 4, April 2004, http://www.jot.fm/issues/issue_2004_04/article8.

[2] J. Bansiya, L. Etzkorn, C. Davis, and W. Li, "A Class Cohesion Metric for Object-Oriented Designs," Journal of Object-Oriented Programming, vol. 11, pp. 47-52, 1999.

[3] C. Blanchi and J. Petrone, "Distributed Interoperable Metadata Registry," *D-Lib Magazine*, vol. 7, no. 12, ISSN 1082-9873, December 2001.

http://www.dlib.org/dlib/december01/blanchi/12blanchi .html.

[4] D. Brutzman, point of contact, *Proceedings of the XMBL/XMSF Workshop on Applicability of the Command & Control Information Exchange Data Model (C2IEDM) for Data Management and Data Interoperability*, Fairfax, VA, March-April, 2004. http://www.vmasc.odu.edu/c2iedm/ws2005.html

[5] M.G. Ceruti, "Mobile Agents in Network-Centric Warfare," *Institute of Electronics Information and Communication Engineers Transactions on Communications,* Tokyo, Japan, vol. E84-B, no. 10, pp. 2781-2785, Oct., 2001.

[6] M.G. Ceruti, "Ontology for Level-One Sensor Fusion and Knowledge Discovery," *Proceedings of the 2004 International Knowledge Discovery and Ontology Workshop (KDO-2004),* 20-24 Sep. 2004.

[7] M.G. Ceruti, and M.N. Kamel, "Preprocessing and Integration of Data from Multiple Sources for Knowledge Discovery," in a special issue of *International Journal on Artificial Intelligence Tools, (IJAIT),* vol. 8, no.2, pp. 152-177, June, 1999.

[8] M.G. Ceruti and B.J. Powers, "Intelligent Agents for FORCEnet: Greater Dependability in Network-Centric Warfare," *Supplemental Volume of the Proceedings of the IEEE International Conference on Dependable Systems and Networks (DSN 2004),* pp. 46-47, June 2004.

[9] M.G. Ceruti and Tedd L. Wright, "Knowledge Management for Distributed-Tracking and the Next-Generation Command and Control," submitted to *11th IEEE International Software Metrics Symposium (METRICS 2005)*, September 2005.

[10] V. Clark, ADM, USN, "Sea Power 21 Series – Part I: Projecting Decisive Joint Capabilities," *Naval Institute Proceedings,* vol. 128, no. 10, pp. 32-41, Oct. 2002.

[11] S. Davis, "Knowledge Engineering: Sensor Ontology," Association of Departments of Computer and Information Science at Minority Institutions Conference, May 2003.

http://webusers.xula.edu/tlopez/SHAUNTRELLE.htm

[12] H. Gao, "Formal Information Fusion Framework (FIFF)" Dec. 2000.

http://www.ece.neu.edu/groups/ifg/demo.ppt

[13] J. Hendler, DAML Sensor Ontology http://www.mindswap.org/~evren/services/sensor-jpa.daml

[14] T. Hengl, "Protégé, Ontology and Knowledge Acquisition: Knowledge Representation, the Foundation of Intelligent Systems," PC-AI Magazine, vol. 18, no. 9, pp. 37-49, 2004.

[15] C.J. Matheus, D. Tribble, M.M. Kokar, M.G. Ceruti and S.C. McGirr, "Towards a Formal Pedigree Ontology for Level-One Sensor Fusion," *Proceedings of the 10th International Command and Control Research and Technology Symposium*, June, 2005. In press.

[16] D. McDaniel, "Multi-Hypothesis Database for Large-Scale Data Fusion," *Proceedings of the 5th International Conference on Information Fusion,* 2002.

[17] D. McDaniel, "Real-time DBMS for Data Fusion," *Proceedings of the 6th International Conference on Information Fusion,* 2003.

[18] S.C. McGirr, "Knowledge Discovery in a Network," *Proceedings of the National Symposium on Sensors and Data Fusion (NSSDF)*, Monterey, CA, May 17-19, 2005.

[19] S.C. McGirr, "Resources for the Design of Data Fusion Systems," *Proceedings of the ISIF Fusion 2001 Conference on Information Fusion*, Montreal, Canada, Aug 10, 2001.

[20] D.L. McGuinness, R. Fikes, J. Hendler and L.A. Stein, "DAML+OWL An Ontology Language for the Semantic Web," *IEEE Intelligent Systems,* vol. 17, no. 5, Sep./Oct. 2002.

[21] D. McGuinness and F. van Harmelen, "OWL Web Ontology Language Overview," 2003, http://www.w3.org/TR/owl-fetures/

[22] NAVMSMO, http://navmsmo.hq.navy.mil/nmsiscat/

[22] J.D. Neushul and Marion G. Ceruti, "Sensor Data Access and Integration Using XML Schemas for FORCEnet," *Space and Naval Warfare Systems Center San Diego Biennial Review, 2005,* In press.

[24] S.L. Reed and D.B. Lenat, "Mapping Ontologies into Cyc," *Proceedings of the AAAI 2002 Conference Workshop on Ontologies for the Semantic Web*, Edmonton, Canada, July 2002

[25] E. Waltz and J. Llinas, *Multisensor Data Fusion,* Artech House, Boston 1990.